Journal of Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Bio-inspired classifier for road extraction from remote sensing imagery

Jiawei Xu
Ruisheng Wang
Shigang Yue

SPIE.

Bio-inspired classifier for road extraction from remote sensing imagery

Jiawei Xu,[a] Ruisheng Wang,[b,*] and Shigang Yue[a]
[a]University of Lincoln, School of Computer Science, Brayford Pool, Lincoln LN6 7TS, United Kingdom
[b]University of Calgary, Department of Geomatics Engineering, Schulich School of Engineering, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada

Abstract. An adaptive approach for road extraction inspired by the mechanism of primary visual cortex (V1) is proposed. The motivation is originated by the characteristics in the receptive field from V1. It has been proved that human or primate visual systems can distinguish useful cues from real scenes effortlessly while traditional computer vision techniques cannot accomplish this task easily. This idea motivates us to design a bio-inspired model for road extraction from remote sensing imagery. The proposed approach is an improved support vector machine (SVM) based on the pooling of feature vectors, using an improved Gaussian radial basis function (RBF) kernel with tuning on synaptic gains. The synaptic gains comprise the feature vectors through an iterative optimization process representing the strength and width of Gaussian RBF kernel. The synaptic gains integrate the excitation and inhibition stimuli based on internal connections from V1. The summation of synaptic gains contributes to pooling of feature vectors. The experimental results verify the correlation between the synaptic gain and classification rules, and then show better performance in comparison with hidden Markov model, SVM, and fuzzy classification approaches. Our contribution is an automatic approach to road extraction without prelabeling and postprocessing work. Another apparent advantage is that our method is robust for images taken even under complex weather conditions such as snowy and foggy weather. [DOI: 10.1117/1.JRS.8.083577]

Keywords: excitation and inhibition; feature vectors; road classification; complex weather.

Paper 14236 received Apr. 22, 2014; revised manuscript received Jun. 4, 2014; accepted for publication Jun. 17, 2014; published online Aug. 5, 2014.

1 Introduction

The population of humanoid robotics such as the miniature air vehicle, unmanned aerial vehicle, and autonomous ground machine increases the demand for outdoor tasks in unknown environments. These tasks include agricultural applications, searching, rescue, as well as military missions. It is necessary that all the applications are required to traverse a variety of terrains. To enable safe and efficient traversal, a module that can correctly obtain the terrain information from given remote sensing images is required. Previous studies can be roughly classified into two categories. The first category is based on proprioceptive sensors.[1–3] The other one is mounting vision sensors on moving vehicles or satellites to obtain remote sensing images.[4–6] Proprioceptive sensors sense the position, the orientation, and the speed of the moving vehicles and provide the navigation service for drivers. Unlike proprioceptive sensors, which are often associated with an inertial measurement unit, measurements from vision sensors are essentially independent of the vehicle speed. However, vision sensors can lead to misclassifications when the ground has a superficial covering or when the environment has reduced visibility due to fog, snow, or other precipitations. In other approaches such as Refs. 7–9, traversable terrains are indicated by the geometry and color from stereo cameras and infrared (IR) range sensors, respectively. These approaches are effective; however, they need much preparation work and high budget. Recently,

*Address all correspondence to: Ruisheng Wang, E-mail: ruiswang@ucalgary.ca
0091-3286/2014/$25.00

machine learning techniques are employed to train the samples and classify remote sensing images. These methods such as support vector machine (SVM) and artificial neural network,[10–12] however, cannot always perform properly and need tremendous presetting and postprocessing effort. Furthermore, to obtain satisfactory results, image processing techniques are often employed to remove "speckle noises" or "residue" in the images. Recent studies, on the other hand, found that human or primate vision systems can extract visual information efficiently and robustly.[13] There have been many articles that analyze the relationship between populations of excitatory and inhibitory connections from human visual cortex.[14–17] The mechanism in primary visual cortex (V1) provides useful methodology to recognize objects under complex environments, which inspires us to develop new methods for road extraction, especially under complex weather conditions. The advantages of our method include nonrequirement of postprocessing. Unlike traditional SVM approach, our method only needs to set training parameters once and therefore is more efficient. In brief, we first collect and normalize corresponding samples of objects to be classified as inhibition or excitation stimuli. Second, the synaptic gain is synchronized by the summation of excitation and inhibition potential on each cell coupling. Third, an improved SVM is employed to recognize roads in remote sensing images by classifying the feature vectors.

The rest of the paper is organized as follows. Section 2 presents the general classification methodology and mathematical formulation. Section 3 further reveals the algorithm and correlation between the classification framework and mechanism from V1. The experimental results and performance evaluation are extensively studied in Sec. 4. Also, we discuss and compare with other state-of-the-art models by considering the noise corruption on remote sensing images. Finally, Sec. 5 concludes the paper and describes future work.

2 Related Work and Methodology

In this section, we propose a SVM-based framework using the improved Gaussian radial basis function (RBF) kernel for remote sensing image classification. First, the synaptic gains sum up the contributions from excitatory and inhibitory stimuli. Second, the synaptic gains compose the feature vectors through an iterative optimization procedure. This procedure filters the synaptic gains and removes the outliers by setting the initialized threshold. Third, a SVM-based classifier using RBF kernel is introduced to segment the images into each specific object (i.e., roads). The background knowledge is given in Sec. 2.1, whereas Sec. 2.2 describes our methodology and model formulation.

2.1 *Related Work*

There are two stages in our framework. The first is the connection of inhibition and excitation stimuli, which yields the synaptic gains for the kernel function and feature vectors. The second is the road recognition by using an improved SVM. Here, we briefly explain the previous work on the inspired mechanism from V1 and SVM as shown in Fig. 1.

2.1.1 *Effects of excitation and inhibition*

Neurophysiological investigations of color stimuli in primary visual cortex (V1) are sensitive to pure color signals and selective for colored patterns.[18] Other irrelevant colored patterns are suppressed according to the visual attention theory. There exists balance of excitation and inhibition in the feedback pathways of visual selection. We notice that, in the context of neurophysiology, balance of excitation and inhibition refers to the relative contributions of excitatory and inhibitory synaptic gains corresponding to some neuronal events, such as oscillation or response evoked by sensory stimulation.[17] This balance indicates that the interplay between excitation and inhibition plays an important role in determining the visual recognition.[2] The understanding of the relationships between these two opposing forces has progressed significantly in the recent years, mainly due to the development of bio-inspired computer vision and pattern recognition techniques.[19]

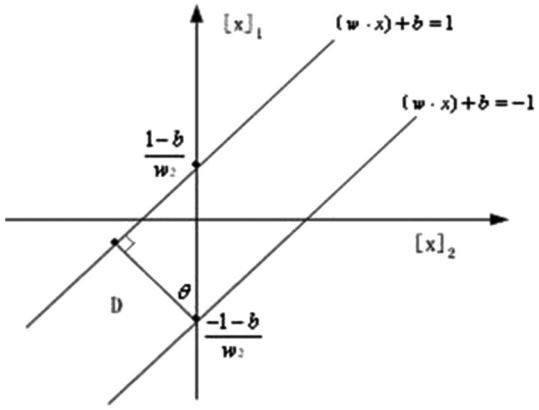

Fig. 1 The interval computation in the support vector machine (SVM).

Our motivation is inspired by the balance between the excitation and inhibition neurons tuning on synaptic gains. By computing contributions of inhibition and excitation, we can obtain a group of synaptic gains, each of which contains properties of a cell's sensitivity to different colors. For a certain color block, the interaction between the excitation and inhibition stimuli corresponds to similar synaptic gains to visual perception. The feature vectors are filtered and selected by setting initialized threshold on synaptic gains. The classifier framework is under an SVM with a Gaussian RBF kernel.[20]

2.1.2 *Support vector machine*

The simplest case of SVM is depicted in Fig. 2 considering only the distance of extreme case in two-dimensional (2-D) space, and the geometric relationship should be

$$\frac{\sqrt{\left(\frac{1-b}{w_1}-\frac{-1-b}{w_1}\right)^2}}{D}=\left|\frac{w_1}{w_2}\right|. \tag{1}$$

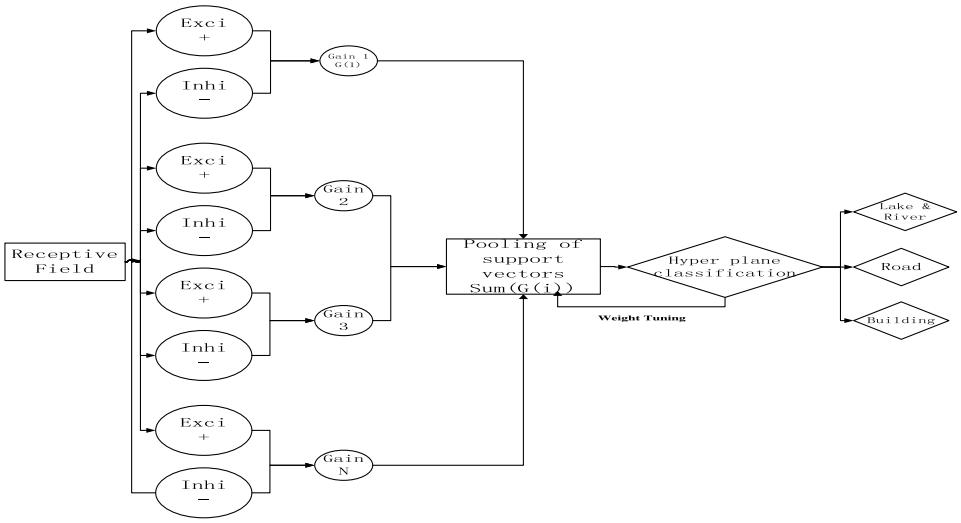

Fig. 2 A schematic illustration of the synaptic coupling on the inhibitory and excitatory neuron cells designed for SVM classification.

Then the distance between the two straight lines is

$$D = \frac{2}{\sqrt{w_1^2 + w_2^2}} = \frac{2}{w}. \tag{2}$$

Suppose that we have the separate line of $\langle w, b\rangle + b = 0$ and the distance of $2/\|w\|$, where $\langle\ ,\ \rangle$ is the inner product. If there exists no hyperplane that can split the "yes" and "no" examples, then the soft margin method will choose a hyperplane that splits the examples as cleanly as possible, while still maximizing the distance between the nearest cleanly split examples

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i \quad \text{subject to } y_i[w^T \phi(x_i) + b] \geq 1 - \xi, \qquad i = 1, \ldots, N, \tag{3}$$

where $\phi(\ )$ is a mapping of low-dimensional vector to one in a higher-dimensional space or feature space, is a penalty, called slack variable, which is introduced for misclassified data, and $C > 0$ is a parameter for increasing or decreasing the penalty factor for classification error. In the case of selecting the coefficients of classifier, from the empirical study, C varies from 2 to 512. The original SVM for remote sensing images classification is proposed by Bruzzone et al.,[21] Inglada,[22] Mountrakis et al.,[23] and Weiss et al.[24] All of them use a set of input data to predict, for each given input, which of two possible classes forms the output. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. Traditional SVM is a supervised statistical learning technique that analyzes data attributes and finds a hyperplane that separates the dataset into multiple classes in a pattern consistent with the training examples. In Sec. 2.2, we analyze our classification model via the improved kernel inspired by the visual mechanism.

2.2 *Proposed Model and Formulation*

The framework of our proposed model is illustrated in Fig. 2. It is proved in biology that the visual recognition is correlated with the competition between excitation and inhibition stimuli. The integration of excitation and inhibition stimuli decides the human vision perception. In biology, this integration generates the synaptic neural gain underlying the perception mechanism in human visual cortex. In other words, assume that we see a scene and the scene comprises an image in our eye. We can sense the different colors in the image because each cell in our eye is sensitive to different colors. If a cell that is sensitive to red color senses the red color, this cell will be the excitation. In the meantime, other cells are suppressed, which means inhibition. Excitatory and inhibitory coupling can have direct antieffect on synchronous neuronal discharge, both of which will jointly decide the final output based on the winner-take-all principle[25] at a certain moment. The synaptic gains generated from this synchronization increase the accuracy and speed of neuronal response. Each of these gains represents a specific block in the image and can be used as a sample for training process, which is an improvement in comparison to the traditional training sample of color features. All synaptic gains are computed from excitation and inhibition potentials, containing singular value and outliers, which can be filtered out by feature vectors through a thresholding process. The pooling of feature vectors is further classified by SVM, and the classifier will output the best weight for each iteration. A detailed description is given in following paragraphs.

The function of receptive field is to receive the color signal within a fixed spatial range. The receptive field can be regarded as the vision sensors or human/primate eyes, while each simple cell in the receptive field is corresponding to each image pixel. Each cell retains the status of either being excitatory or being inhibitory. The excitation and inhibition nodes indicate the integration of the interested (excited) object (e.g., roads) and suppressed object (e.g., nonroad), respectively, which generate synaptic gain $G(i)$ $(1,2,\ldots,N)$. These synaptic gains contribute to the pooling of feature vectors and we adopt SVM for the image classification. Feature vectors form a collection by analyzing the thresholds in the synaptic gains. The last step is to separate these vectors into several classes for each specific object.

The formulation of our framework is elaborated in the following. The receptive fields of ganglion cells innervated by cones have a center-surrounding organization. However, in the case of the cone pathway, the center and the surrounding may be color opponent. This means that the ganglion cell is excited by a specific color stimulus (one object class) in the center of the receptive field and inhibited by the "opposite" color (other object classes) in the surrounding area. Each cell corresponds to one image pixel in the input images. First, we adopt the multilevel segmentation to obtain an object-based decomposition of remote sensing images,[26] while Table 1 indicates our improvement on this example learning and image segmentation processing techniques. Then, the excitation and inhibition classes can be decomposed into the interesting and uninteresting objects in each image such as roads or lakes. It can be approximately estimated for the excitation contribution on each node with the following mathematical formulation

$$G(\text{Excitation}) = \frac{|1 + m_x|}{\sum_{x=1} m_x}, \tag{4}$$

where m_x is the corresponding segmented pixel blocks in color values by applying multilevel segmentation ($x = 1,2,\ldots,N$).[26]

Accordingly, the inhibition contribution in primary visual cortex can also be quantitatively recorded using Eq. (2).

$$G(\text{Inhibition}) = \frac{|1 - m_x|}{\sum_{x=1} m_x}. \tag{5}$$

To quantitatively define the relationship of the excitation and inhibition neural activity, we need to consider the potential loss during the transmission. The summation excitation and inhibition is usually less than 1 for each coupling. Physically, the synaptic gain is the computational integration of each node's excitation and inhibition, which is the most likely of each segmented image block. Therefore, the synaptic gain $G(i)$ can be computed as the summation of $G(\text{Excitation})$ and $G(\text{Inhibition})$ in primary visual cortex regarding one node i. If a cell i ($i \neq j$) discharges at the node j, the function $G(i)$ is augmented as

$$G(i) \to 1 + G(j)/G(i), \tag{6}$$

where the updated $G(i)$ is the contribution from one synaptic coupling by importing new synaptic nodes. We consider the summation of all synaptic gains by computing the following equation:

$$\sum_{i=1} G(i) = G(i) * \beta + \alpha, \tag{7}$$

where α and β are empirical parameters tuning the strength and kernel width of the synaptic gains, respectively. Then, we can get $\sum_{i=1} G(i)$ with the increment of node $G(i)$ and map them onto the pooling of feature vectors after filtering the outliers by setting the empirical threshold β. This threshold is a nonlinear one imposed by the spike-rate-encoding mechanism ranging from 0 to 1. Thereby we initialize β at 0.614 as the default value for lake and river, 0.57 for road,

Table 1 Overall classification accuracy (%) by comparing the proposed approach and the approach of Xua and Duygulub.[26]

	True positive	True negative	False positive	False negative
Object-based image segmentation	4.1	4.5	65.3	25.1
Proposed method	6.8	9.4	61.4	22.4

Table 2 Results by cell nodes of $G(i)$ for lake terrain.

Nodes of $G(i)$	Accuracy (%)	Feature vectors
190	65.37	117
253	69.52	98
775	**76.79**	**382**
917	64.78	82
1214	62.11%	92

Note: Bold values indicate best performance values.

Table 3 Optimized parameters tuning on α and β for radial basis function (RBF) kernel function.

Terrain	α	β	Minimum error rate (%)
Lake/river	17.2	0.679	23.21
Road	23.8	0.521	17.82
Building	20.7	0.626	21.17

and 0.595 for building terrain. These parameters will be optimized with the incremental synaptic gains $\sum_{i=1} G(i)$.

To classify effective synaptic gains (i.e., feature vectors) in the hyperplane space, several kernels have been proposed such as linear kernel, polynomial kernel, RBF kernel, and sigmoid kernel. By comparing with other kernel functions, we finally select a transformed RBF kernel for the feature vectors due to the best accuracy result (see Tables 2 and 3), which is expressed as follows:

Improved RBF kernel

$$G(x) = \exp\left(-\frac{|G(i) - \overline{G(i)}|^2}{\alpha^2 \beta}\right), \tag{8}$$

where $G(x)$ is unknown synaptic gains to be classified and $\overline{G(i)}$ is the average potential of synaptic gain. It then projects the samples to higher dimension using the following kernel function as described below. Kernels can be regarded as a generalized inner product.

In this improved RBF kernel, each center of radial function corresponds to one feature vector. Tuning α and β from RBF kernel, we can classify the query samples into a certain class with the maximum summation of the absolute values of coefficients. Here, the decision rule is expressed as

$$\text{Identity}(y) = \text{Class}(\max_{i=1,2,\ldots,N} \|G(x)\|), \tag{9}$$

where N denotes the number of the feature vector in segmented block i. Class indicates the classified feature vectors comprising the outputs. Finally, the feature vector classifiers automatically recognize the objects represented by pixel blocks and stop iterations until all the pixel blocks are classified into corresponding objects or background.

3 Analysis of Bio-Inspired Classifier

This section analyzes the basic flowchart on the excitation/inhibition classifier framework. The final aim is to tune the parameters for RBF kernel, which decompose the problem into the training sample collection and optimize two subproblems. Table 4 indicates the algorithm inspired by the mechanism from V1.

Table 4 The bio-inspired training and classification algorithm.

1. Adopting the multi-level segmentation to obtain an object-based decomposition of the image.

2. Collecting training samples from several remote sensing images and combining the data into the documents, with interested terrain (one object class) and irrelevant terrains (other object classes) saved as 'excitation.txt' and 'inhibition.txt'.

3. **Input**: A Training sample (prototype) matrix $G(i) = [G(1), G(2), \ldots, G(N)] \in R^{n*N}$, a query sample

4. Computing $G(i)$ with the increment on neural nodes in Eq. (3).

5. Tuning α and β for RBF kernel to obtain synaptic gain $\sum G(i)$ from Eq. (4).

6. Filtering and removing the outliers from synaptic gain to form feature vectors pooling.

7. Searching the hyper-plane from improved support vector machine using various kernel functions and choosing the optimized kernel function

8. **Output:** $\text{Identity}(y) = \text{Class}(\max_{i=1,2,\ldots,N} \|G(x)\|)$

9. Converting the classified matrix into images, setting the irrelevant data as the background.

Before analyzing the relationship and effect of the excitatory and inhibitory cells, the prerequisite is to segment the original input images to the rough blocks while keeping the miniatures of the input information. We adopt the multilevel segmentation to obtain an object-based decomposition of the image.[26] Several sample blocks are collected from each object and then converted into the excitation/inhibition stimuli. The summations of each coupling (excitation/inhibition) are accumulated to formalize the pooling of feature vectors. Then the process turns into SVM classification problems shown in Table 4. Figure 3 depicts the intuitive explanation of our framework and implementation of the algorithm. Blocks in different colors contain the corresponding excitation/inhibition coupling neurons. We trained three groups of object samples in the learning process: roads, lakes, and buildings. Other remaining objects are classified into "others" (see Table 5).

The training samples are collected from 15 remote sensing images. The original image is first converted into the matrix form, and the corresponding object is cropped from the image to serve as training sample. Then we stitch the patches for a certain sample class (e.g., lakes, roads, or buildings) together and save them in a text document. For example, if the interested object is road, the road region is collected as the excitatory samples, while the others are saved as inhibitory data as shown in Fig. 4. The red region indicates a segmented block that corresponds to a specific object in the original image for excited neural cell block. The blue rectangles indicate the segmented blocks with other irrelevant objects for suppressed neural cell blocks.

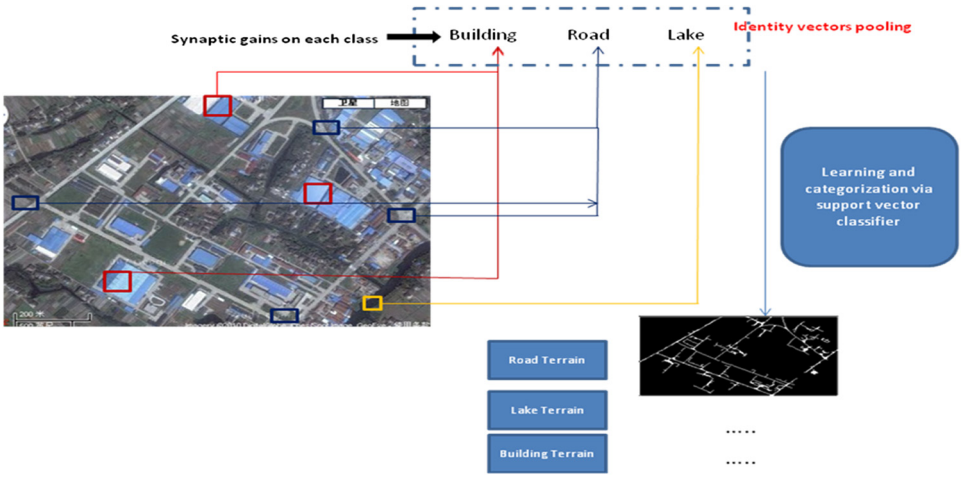

Fig. 3 An intuitive explanation of the training and classification of the proposed feature vector classifier.

Table 5 Classification accuracy of the proposed algorithm with reference by ground truth.

	Lake/river	Road	Building	Others
Lake/river	**89.6**	1.2	5.7	3.5
Road	3.9	**85.6**	7.8	2.7
Building	5.3	3.1	**87.2**	4.4

	53	54	55	56	57	58	59	60	61	62
22	192	184	200	203	194	172	146	188	192	183
23	180	169	138	130	137	140	144	200	205	188
24	176	162	126	103	98	98	123	168	173	188
25	84	83	75	76	85	94	134	168	172	193
26	38	38	41	45	50	79	116	138	157	186
27	44	45	60	63	62	74	106	115	139	177
28	62	73	102	95	75	77	106	108	126	164
29	67	74	109	94	79	84	107	107	106	146
30	86	84	120	106	88	90	107	108	104	150
31	113	90	104	113	100	96	97	99	104	140
32	114	88	88	106	96	113	116	120	126	121
33	110	87	93	104	98	121	124	143	147	114
34	117	89	105	121	122	141	136	152	144	106
35	151	139	148	146	145	159	158	174	158	113
36	176	173	169	164	157	156	155	157	133	99
37	182	182	174	175	159	150	144	132	119	96
38	183	181	174	175	155	150	139	123	118	110
39	178	178	173	166	155	154	145	134	123	109
40	166	164	162	163	164	161	154	152	144	129
41	167	164	160	165	171	169	145	136	134	131

Fig. 4 An example of how we collect the excitation and inhibition stimuli for a specific terrain, where the excitation and inhibition values come from these prelearning stimuli. (Red for 'excitation.txt' and others for 'inhibition.txt').

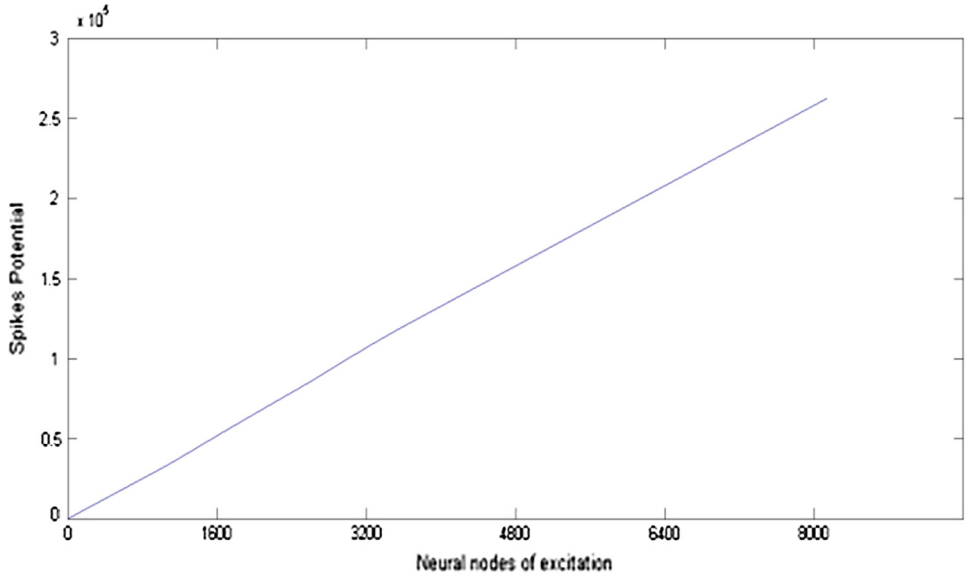

Fig. 5 The relationship between the potential and length of training samples. This figure indicates the excitation stimuli. *x*-axis denotes the neural nodes and *y*-axis indicates the neural spikes potential.

Excitation and inhibition cells contribute the synaptic gains which are determined by the effect of internal counterbalance. Figures 5 and 6 show the relationship between excitation and inhibition initialization function. The horizontal axes are recorded with the synchronization and the vertical axes are normalized to the same extent.

In Fig. 7, it shows that the synaptic gain of the neuron nodes is proportional to the increment of feature vectors. There are two assumptions for this property. First, we assume that the complex cells are contrast-invariant to excitatory and inhibitory simple cells. Second, the relative peak adjustment of membrane potential fluctuations between individual cells is stable. Based on this assumption, the relationship between the synaptic gains and feature vectors is plotted as shown in Figs. 5 and 6. The vertical axis represents the total synaptic gains in the receptive field, and the horizontal axis indicates the feature vectors by optimizing the β. In general, the mortality rate is about 2:1, which means that half of the synaptic gains are selected for the feature vectors.

In Fig. 8 and Table 2, the illustration records the total number of feature vectors for the lake object in one test image. In Fig. 8, the amount of the best feature vectors is 382 for the classification of this specific object. The x label indicates the amount of feature vectors,

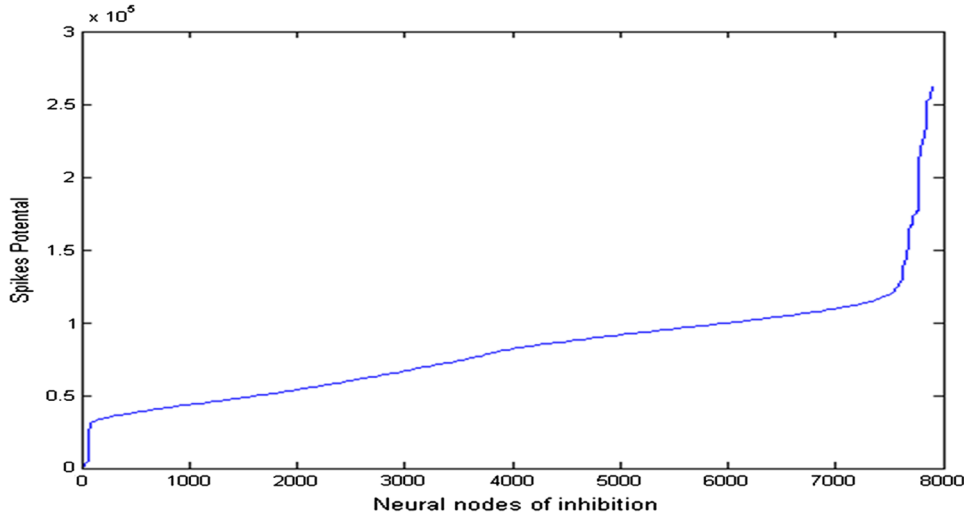

Fig. 6 The relationship between the potential and length of training samples. This figure indicates the excitation stimuli. x-axis denotes the neural nodes and y-axis indicates the neural spikes potential.

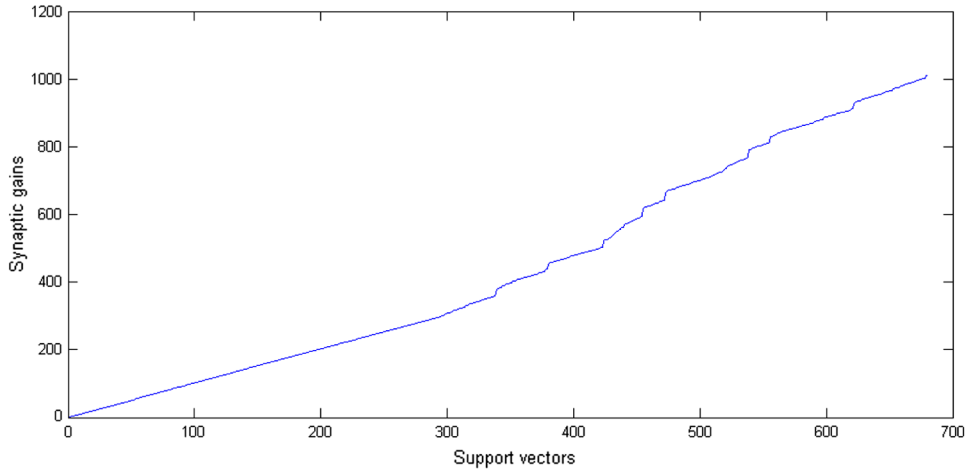

Fig. 7 The relationship between synaptic gain nodes and support vectors.

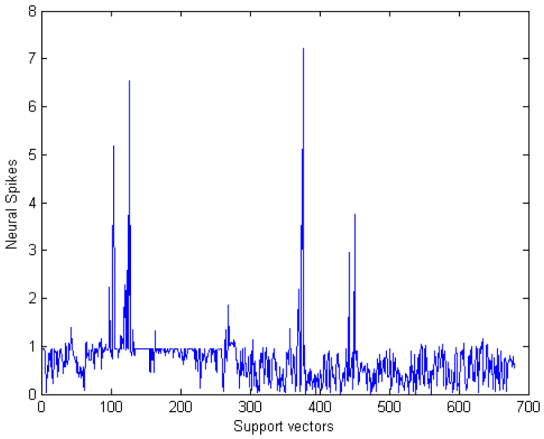

Fig. 8 The spiking curve of feature vectors for the lake terrain in one testing image.

and y label represents the spiking potential which is mostly distinguishable in our classifier framework.

It is found that a relatively larger value will give a smoother decision surface and more regular decision boundary. This is because the RBF kernel with large values allows feature vectors to have strong influence over a larger area. The experimental results in Table 3 further prove this assumption quantitatively. In Table 3, the parameters α and β denote the strength and kernel width of the synaptic gains, respectively. The kernel width is initialized and then updated to find the best output of the classification.

In our experiment, we extensively tested the model to obtain the highest accuracy for three types of object classification. The description of quantitative and qualitative comparisons is given in experimental sections.

4 Experimental Results and Analysis

To demonstrate the effectiveness of the proposed adaptive classification model, our method has been extensively tested using several types of remote sensing images from the benchmarks. The detail of the dataset is given in Table 6. All tests in this section are implemented in MATLAB 2010b and performed on the Windows 7 platform. The experiments further analyze the relationship between synaptic gains and the object classification and also validate the performance by comparing other state-of-the-art methods under normal and difficult weather conditions. Experiments were carried out on a PC (CPU: Intel ® Core ™2 Quad Q8400 2.66 GHz, RAM: 2 GB).

4.1 *Experiment Setup*

The proposed method has been tested with different benchmark databases to verify its feasibility and robustness. The source of the remote sensing images is described in Table 6, including natural images, artificial images, and other video frames. The testing images can be downloaded at Remote Sensing Instrument Database,[27] Remote Sensing Images from University of Essex Databases,[28] Resource and Environment Remote Sensing Database.[29]

Figure 9 shows the sample training process, and all the data are normalized and processed to generate a text file. Figure 10 shows the initialization of excitation and inhibition stimuli. The inhibition stimuli affect the summation of the synaptic gains as we can find the troughs shown in

Table 6 Testing datasets for the proposed model.

Group	Category	Subjects(terrain)	Images
Benchmark database[1,2,3]	Remote sensing Images	Lake & river, road, buildings	65

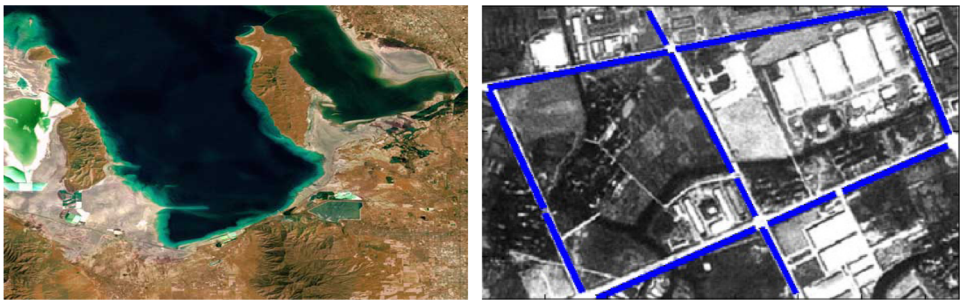

Fig. 9 An example of our training data collected from the benchmark remote sensing images: the left green region shows the lake/river regions and the right one with blue straight lines indicate the road class for training samples, they are converted into matrix for further processing.

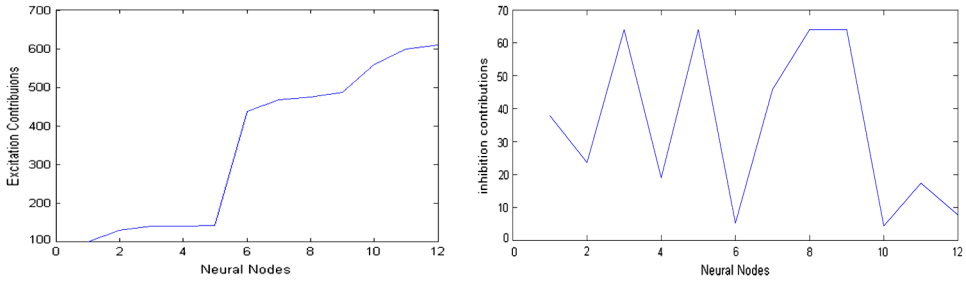

Fig. 10 Illustration of the initialization of the excitation and inhibition stimuli (x label indicates the neuron nodes and y label represents the excitation and inhibition contributions, respectively).

the right figure. These stimuli yield synaptic gains and classifier hyperplane for the feature vector classifier.

4.2 *Experimental Results*

In this section, extensive experiments are tested to validate the performance of the excitation/inhibition classifier using the benchmark datasets under normal weather condition. Data were randomly permuted 10 times, and the standard deviations are shown in Table 7. A typical result is shown in Fig. 11. In this case, we successfully classify the image into roads, lakes/rivers, and others. In Fig. 12, we illustrate the corresponding potential from the testing data for different objects in this remote sensing image, such as roads, lakes, and buildings. The green curve indicates the synaptic gain potential of the lake class on integration by excitation and inhibition stimuli, the blue one represents the building class, and the red one indicates the road class. The potential levels of feature vectors are evolved with the iteration of synaptic gain. As shown in Fig. 12, the potential difference of three objects shows a relatively large deviation, which indicates a relatively distinguishable distribution for the testing samples classification.

Table 7 Standard deviation for each training sample.

Resampling for N time	Standard deviation	Resampling for N time	Standard deviation
1	1.3911	6	0.8647
2	1.2361	7	1.1925
3	1.3731	8	1.2432
4	0.9986	9	1.3056
5	1.2534	10	1.2793

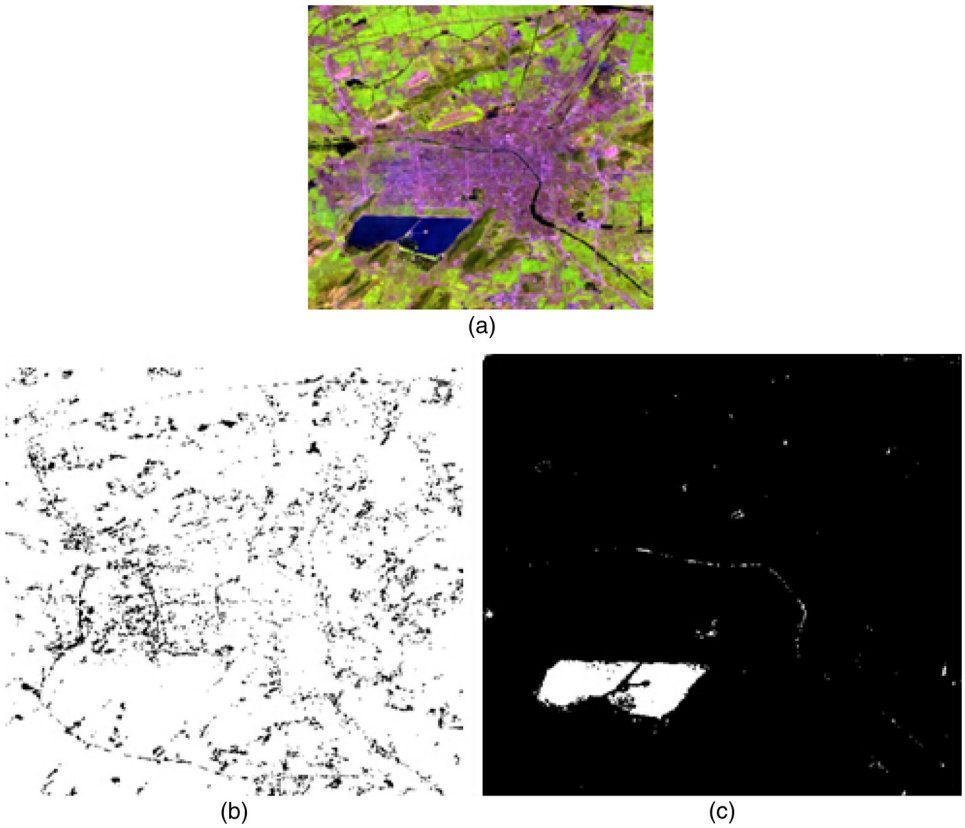

Fig. 11 A typical classification example from high resolution satellite images (image size 300×300), (a) indicates the original image from the benchmark database, (b) represents the classified roads (shown in black), (c) shows the lakes and rivers (shown in white; alternatively we can regard this region as the excitation class, while the black area is suppressed as the inhibition class).

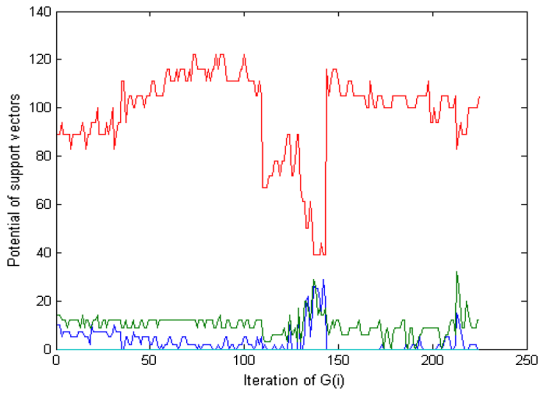

Fig. 12 The neural potential from the classified data in Fig. 9. (The green curve indicates the lake class, the blue curve represents the building class, and the red curve indicates the road class).

Furthermore, in Figs. 13–16, we select some results from the benchmark database and we show the results of the road extraction. Other objects can also be obtained by using the proposed training and testing method. The corresponding action potential is illustrated under each result. The qualitative analysis and comparisons with other classification methods are given in Sec. 4.3. In Sec. 4.4, we discussed a critical issue on the classification under complex weather conditions.

In Fig. 12, we demonstrate four results for road classification. All of them are selected from remote sensing image database: group (a) image is occupied with buildings and roads, geomorphic features between groups (b) and (c) are quite similar, group (d) is a grayscale image. The

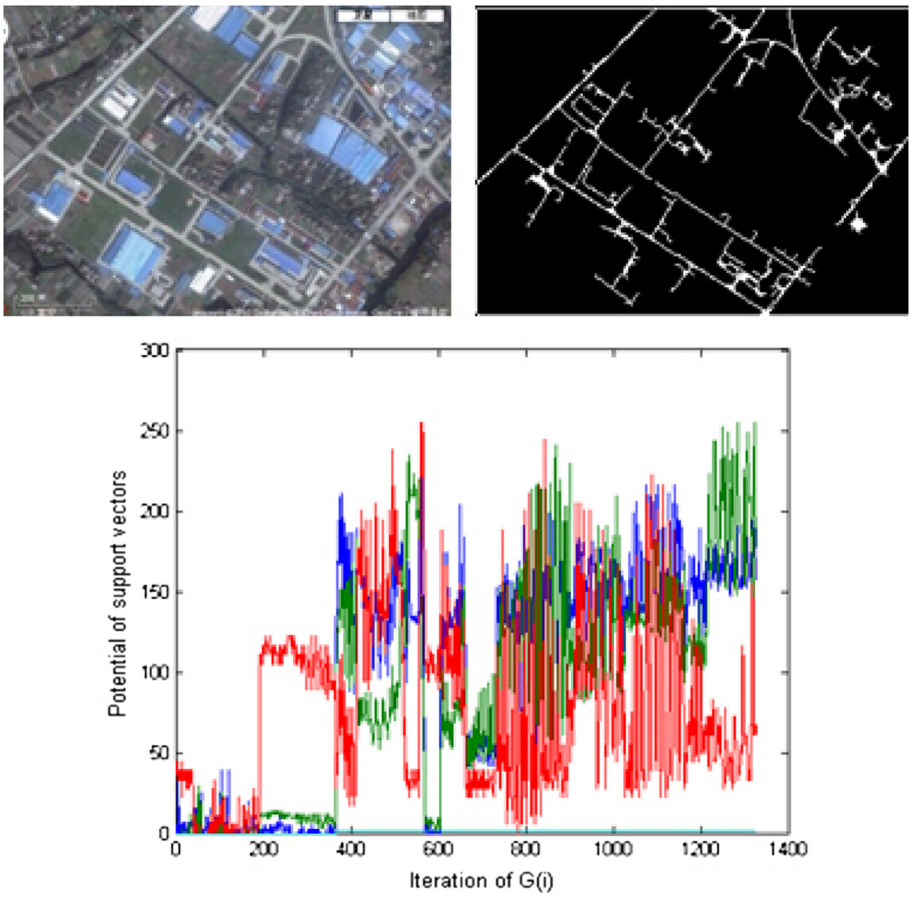

Fig. 13 One experimental result from high-resolution remote sensing images; (a) indicates the original remote sensing image, (b) is the classified road image, (c) of each group is the corresponding synaptic gain regarding the different objects in the corresponding image. The green curve records the potential of feature vectors for the road class with the classified feature vectors, the red curve is the building class, and the blue curve is the grassland class.

roads are classified successfully and we further discuss its performance in Sec. 4.3. All other objects, such as buildings and rivers, are suppressed by the inhibition class from V1, whereas the roads are extracted.

4.3 *Performance Evaluation and Comparison with Other State-of-the-Art Models*

To make quantitative comparisons with the state-of-the-art models, we use ground truth data as reference and set up the equivalent evaluation methods. We first compare three models under the normal weather condition.[4,30,31] Tso and Olsen[30] combined spectral and spatial information into hidden Markov field models for unsupervised image classification; Zhang and Foody[31] proposed fuzzy classification for the remote sensing images. Melgani and Bruzzone[4] applied binary SVMs to multiclass problems in hyperspectral data.

Generally, there are six criteria in the performance evaluation.[32,33] They are percent complete, percent correct, rank distance, quality percentage, branch factor, and percent redundancy, respectively. Rank distance measures the normalized distance (in completeness and correctness space) between the result and the reference (ground truth), and the quality percentage measures the absolute quality of the detected objects on the road. To obtain 100% quality, the ground truth must be correctly labeled to indicate the warning by users, without missing any one ($FN = 0$), and without mislabeling any background pixels ($FP = 0$). For the purpose of evaluation, we have made the reference model of each test frame manually.

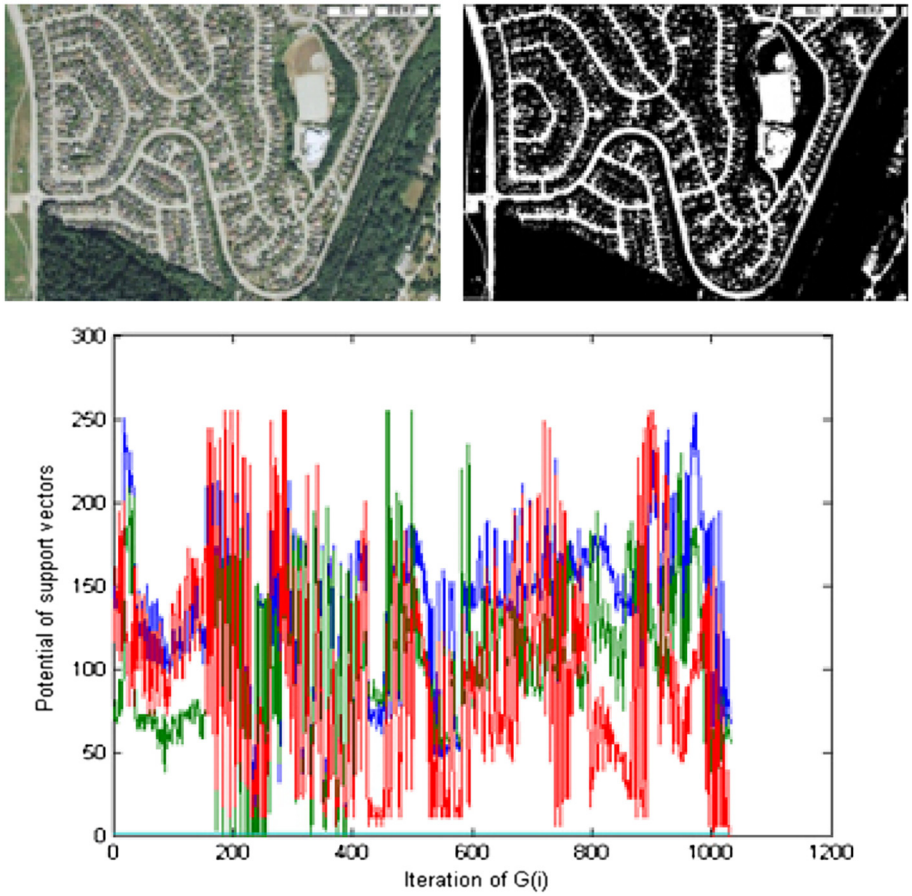

Fig. 14 One experimental result from high-resolution remote sensing images.

Some of the widely used criteria are defined as following:

1. Percent complete:

$$\frac{100\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}. \tag{10}$$

2. Percent correct:

$$\frac{100\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}. \tag{11}$$

3. Rank distance:

$$\sqrt{\frac{\%\mathrm{complete}^2+\%\mathrm{correct}^2}{2}}. \tag{12}$$

4. Quality percentage:

$$\frac{100\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}. \tag{13}$$

5. Branch factor:

$$\frac{\mathrm{FP}}{\mathrm{TP}}. \tag{14}$$

6. Percent redundancy: the percentage of the generated warning output that overlaps itself.

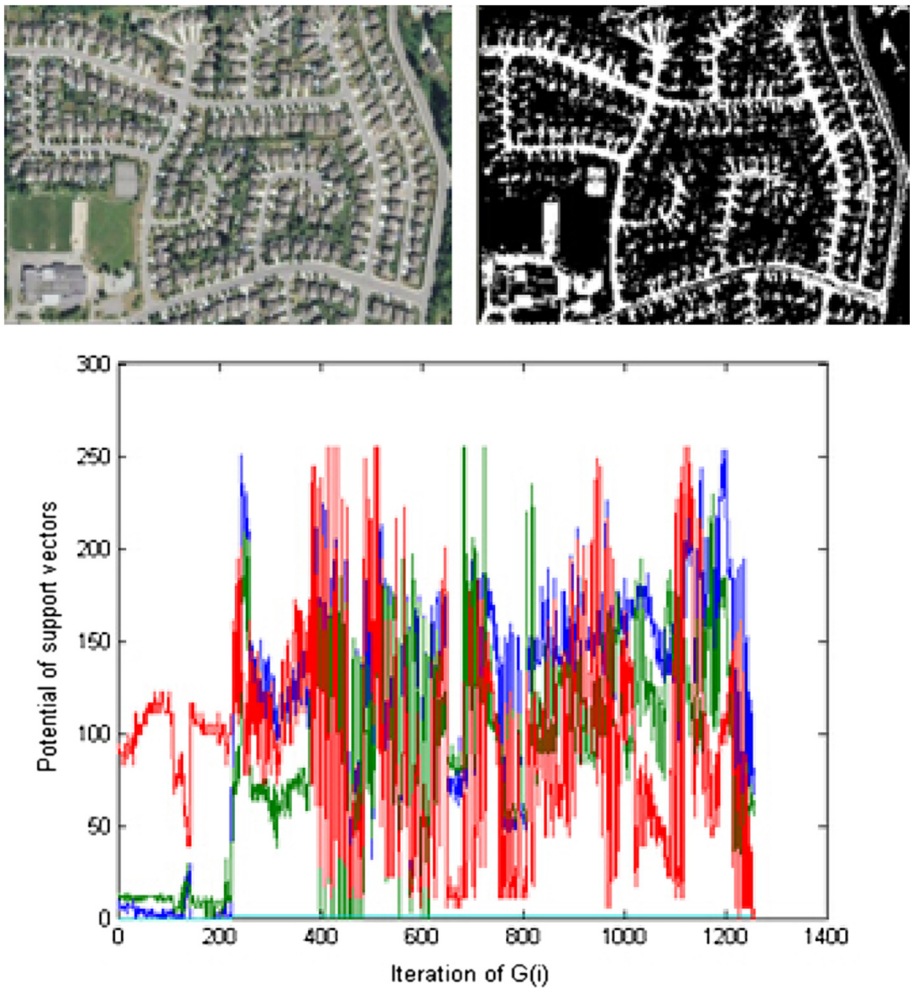

Fig. 15 One experimental result from high-resolution remote sensing images.

The percent complete measures the percentage of the reference model that is covered by the derived model and ranges from 0% to 100%. The percent correct is a similar measure but gives the percentage of the derived model covered by the reference model instead of the other way around, where the high values are best. The rank distance is a new measure of the overall quality of the result. It measures the normalized distance (in completeness and correctness space) between the result and the reference model. Like the other two measures, it ranges from 0% to 100%, where the higher values are better. Finally, the percent redundancy is useful for determining how much extra work is being done. Figure 17 illustrated selected results by our experiments; the top three row images are normalized to 1600×1200 pixels, while the bottom row images indicate the 300×300 pixels. In Tables 5 and 8–10, we compared the classification accuracy with Tso, Zhang, Melgani, and our approach, where our method outperforms other methods under the normal weather condition. In fuzzy classification simulation, we tuned the parameters during the simulation and led the deviation of each specific class in the results listed in the Table 8. The results in Fig. 18 indicate the lower FN, FP, and relatively higher TP, TN, percent complete, which means a better performance on this situation. In Table 11, the average running time is recorded for each terrain. The proposed method is relatively time consuming with the complexity of classification features.

To summarize, the classification performance of excitation/inhibition classifier outperforms Markov random field and fuzzy classification. The experiment results positively reflect the superiority of the bio-inspired mechanism from primary visual cortex. Compared with other state-of-the-art terrain classifiers, such as HMM, fuzzy classification, and traditional SVM, the experiment results also show a better performance under complex weather condition, which will be discussed in the following. Their official website illustrated these models clearly.

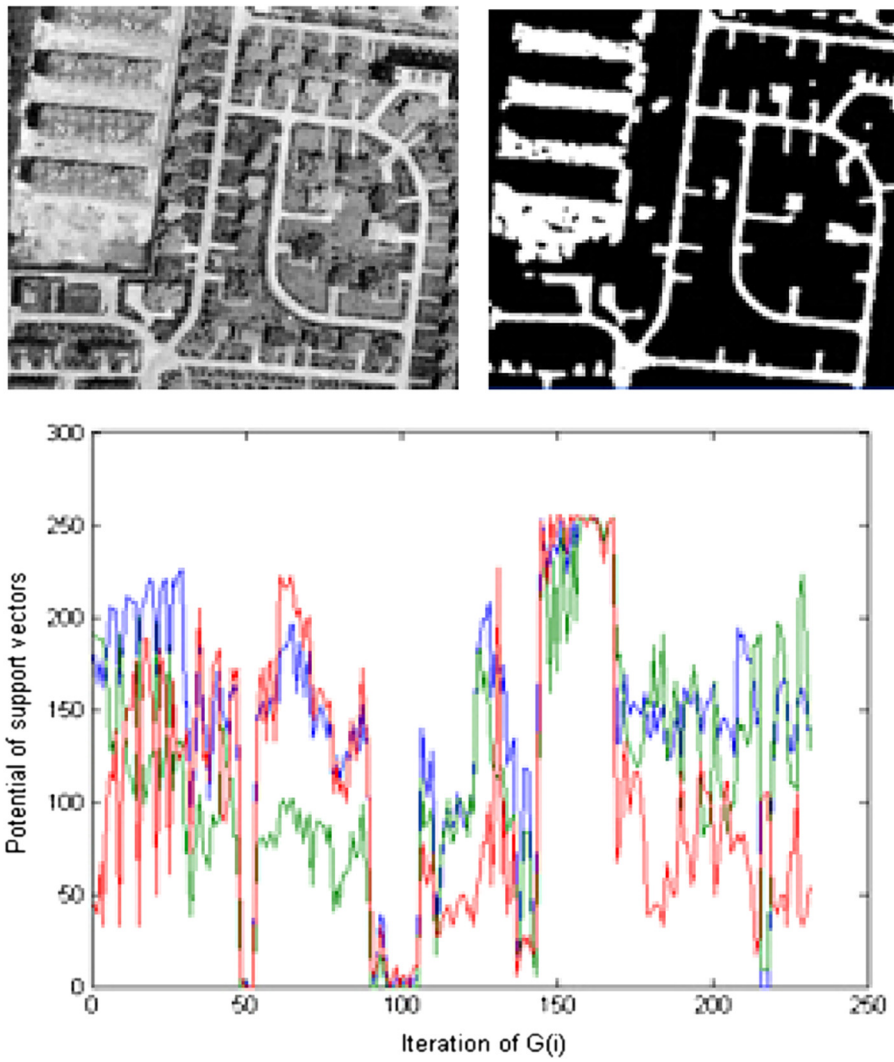

Fig. 16 One experimental result from high-resolution remote sensing images.

For example, HMM is released with C++, while fuzzy classification and SVM are given in MATLAB code.

4.4 *Comparisons under Complex Weather Condition and Discussion*

It is important to show that this adaptive classification method is robust to external noise. Experiments prove that it can reduce the noise effectively using Gabor or median filters,[34,35] as they are good fits to the receptive field in primary visual cortex (V1). We consider images taken under foggy and snowy weather condition as shown in Figs. 19 and 20. The artificial foggy and snowy weather conditions are made on the original image, with corruption ratio from 0% to 50%, interval of 10%. In Figs. 21 and 22, the corresponding results are demonstrated below. The classified road and background are shown in black and white and grayscale, respectively.

The intention of this processing is trying to validate the robustness of our method in comparison to other cutting-edge models under different weather conditions. Tables 12 and 13 reveal the satisfactory performance comparing with other cutting-edge models. A brief discussion is given below to analyze the potential reasons.

From Table 12, we realize that our algorithm outperforms other methods at the foggy ratio 0%, 10%, 30%, and 50%. However, hidden Markov model (HMM) dominates the advantage at the ratio of 20% and 40%. Considering that HMM is based on the simulated annealing algorithm to optimize the objective function, it is relatively robust to the noise at the foggy ratio of 20% and

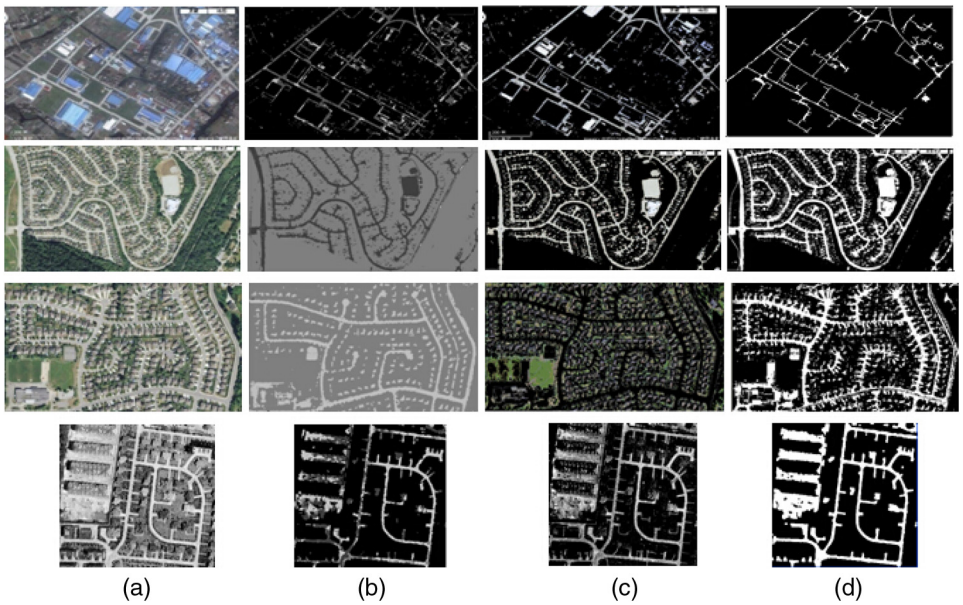

Fig. 17 Column (a) indicates original remote sensing images, column (b) represents results by using hidden Markov random field model, column (c) indicates results by using fuzzy classification, and column (d) illustrates the results from the proposed model.

Table 8 Classification accuracy (%) of hidden Markov model (HMM) with reference by ground truth.

	Lake/river	Road	Building	Others
Lake/river	81.1	9.6	4.3	5.0
Road	10.8	71.3	11.7	6.2
Building	12.7	5.8	76.1	5.4

Table 9 Classification accuracy of fuzzy classification with reference by ground truth.

	Lake/river	Road	Building	Others
Lake/river	75.2 ± 4.3	6.4	11.2	5.2
Road	10.1	79.1 ± 3.9	4.8	6.0
Building	4.1	5.2	83.4 ± 4.8	7.3

Table 10 Classification accuracy of improved support vector machine (SVM) with reference by ground truth.

	Lake/river	Road	Building	Others
Lake/river	**85.7**	4.8	5.3	5.2
Road	4.2	**84.5**	5.8	5.5
Building	5.8	4.1	**84.0**	6.1

Note: Bold values indicate best performance values.

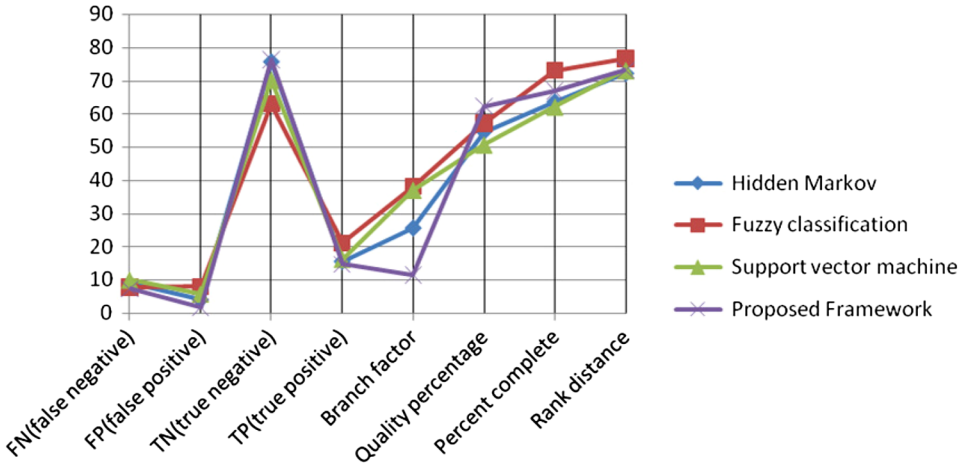

Fig. 18 Performance comparisons with other state-of-the-art models based on the widely used criteria [from Eqs. (10)–(14)].

Table 11 The average running time for each terrain.

	Lake/river	Road	Building
HMM[22]	10.92 s	18.24 s	16.47 s
Fuzzy classification[23]	13.87 s	14.41 s	13.98 s
SVM[10]	15.36 s	19.73 s	18.70 s
Proposed methods	19.57 s	23.65 s	21.43 s

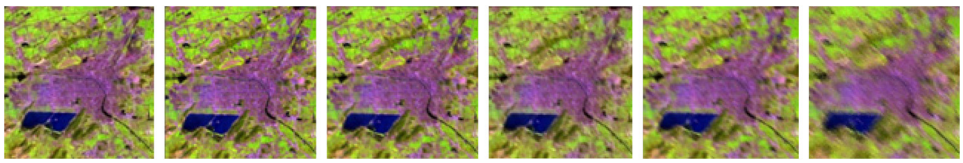

Fig. 19 From left to right, each image illustrates sample under foggy weather ratio from 0% to 50%, interval of 10%.

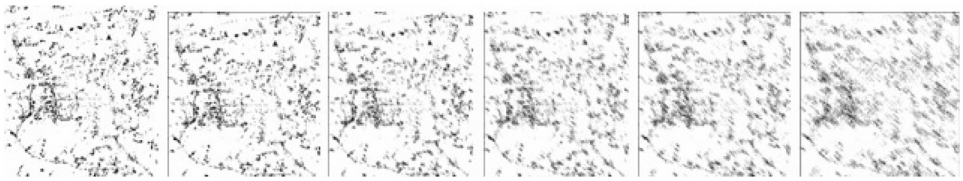

Fig. 20 From left to right, each image illustrates sample under snowy weather ratio from 0% to 50%, interval of 10%.

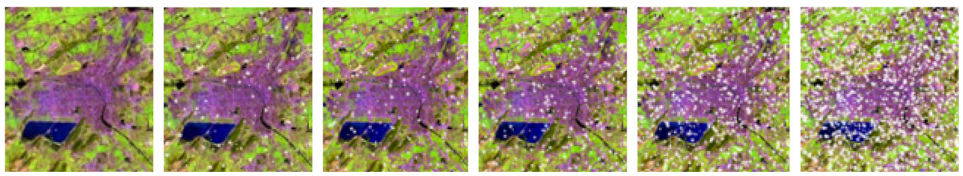

Fig. 21 From left to right, corresponding classified road image under foggy weather ratio from 0% to 50% is illustrated, interval of 10%.

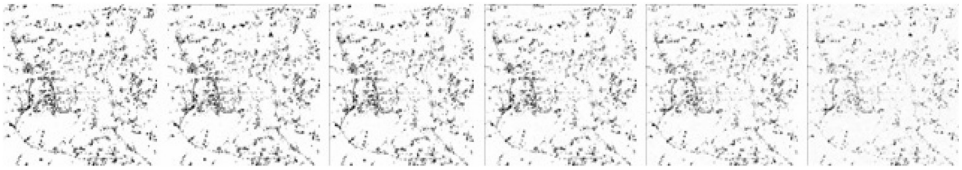

Fig. 22 From left to right, corresponding classified road image under snowy weather ratio from 0% to 50% is illustrated, interval of 10%.

Table 12 The recognition rates (%) under different levels of foggy effects regarding Fig. 16.

	0	10	20	30	40	50
HMM[22]	95.0	85.3	**84.2**	81.7	**77.2**	69.7
Fuzzy classification[23]	94.8	81.7	79.2	78.4	68.2	58.1
SVM[10]	93.8	86.9	80.3	79.9	76.1	68.3
Proposed methods	**96.5**	**87.8**	83.9	**82.7**	76.3	**70.1**

Note: Bold values indicate best performance values.

Table 13 The recognition rates (%) under different levels of snowy effects regarding Fig. 18.

	0	10	20	30	40	50
HMM[22]	95.0	85.3	81.5	64.9	52.3	44.2
Fuzzy classification[23]	94.8	85.7	69.3	56.5	49.2	37.3
SVM[10]	93.8	84.9	**83.1**	66.2	54.5	47.3
Proposed methods	**96.5**	**86.1**	76.9	**67.3**	**57.5**	**48.1**

Note: Bold values indicate best performance values.

40%. Our proposed method has been automatically optimized to obtain the best fitness value to approach the human visual perception mechanism. The classifier can adaptively find the optimal hyperplane for the new test query, which enhances the performance and robustness under the noise-corrupted image.

In Table 13, although traditional SVM gives its advantage at snowy corruption ratio 20%, the proposed method shows excellent performance under this snowy weather. Especially at a higher corruption ratio of 40% and 50%, the proposed approach is nearly 10% better than the other approaches. In most cases, our method is still superior to fuzzy classification, HMM, and traditional SVM. The experiment results provide strong proofs for our analysis in Secs. 3 and 4. It lays a biologically inspired foundation for remote sensing image classification under complex weather conditions.

5 Conclusions

In this paper, we introduce a bio-inspired method based on the visual perception mechanism to classify remote sensing images and compute the synaptic gains by summing up the excitation and inhibition stimuli. It optimizes the parameters by tuning the strength and width of kernel function for the query samples. This approach is an automatic learning-detection method which relieves the manual work on postprocessing from the remote sensing images.

Moreover, another merit is the robustness under complex external noise, such as foggy or snowy weather as shown in Sec. 4.4. This is an important contribution, as objects in optical images are not always imaged properly due to different weather conditions. Traditional classification methods are frequently influenced by severe weather conditions. In the future, we will

further validate the bio-inspired visual model on the application of pattern recognition problems, such as objects detection, image or video categorization, and so forth.

To summarize, this paper proposes an early biologically inspired framework for traditional classification problems for remote sensing images. There are also some limitations in the proposed method. For example, one drawback is that our model needs more CPU time than fuzzy clustering method and cannot fundamentally implement the algorithm for real-time systems. We will continue to explore more robust bio-inspired computational models with visual features in the future. In addition, previous studies[36,37] addressed that motion cues frequently dominate the visual decision under complex world. This critical factor will be further integrated to the proposed framework in the future work.

Acknowledgments

The authors thank all of the collaborators whose modeling work is reviewed here and to the members of School of Computer Science at the University of Lincoln for discussion and feedback on this research. This work was supported in part by grants of EU FP7-IRSES Project EYE2E (Building a Visual Brain for Fast Human Machine Interaction, 269118), LIVCODE (Life like Information Processing for Robust Collision Detection, 295151), and HAZCEPT (Nature Inspired Hazard Perception—towards Zero Road Accidents, 318907).

References

1. A. Angelova et al., "Fast terrain classification using variable-length representation for autonomous navigation," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, Piscataway, NJ, USA, (2007).
2. M. R. DeWeese and A. M. Zador, "Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex," *J. Neurosci.* **26**, 12206–12218 (2006).
3. V. Pedroni, "Inhibitory mechanism analysis of complexity O(N) MOS winner-take-all networks," *IEEE Trans. Circuits Syst.* **42**(3), 172–175 (1995).
4. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.* **42**(8), 1778–1790 (2004).
5. A. Rankin and L. Matthies, "Daytime mud detection for unmanned ground vehicle autonomous navigation," in *Proc. 26th Army Science Conf.*, California Institute of Technology Pasadena Jet Propulsion Lab (2008).
6. D. Sadhukhan and C. Moore, "Online terrain estimation using internal sensors," in *Proc. Florida Conf. Recent Advances in Robotics*, Boca Raton, FL (2003).
7. L. Bruzzone, D. Prieto, and S. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.* **37**(3), 1350–1359 (1999).
8. H. Heinze et al., "Visual event-related potentials index focused attention within bilateral arrays. I. Evidence for early selection," *Electroencephalograph. Clin. Neurophysiol.* **75**(6), 528–542 (1990).
9. C. A. Brooks and K. Iagnemma, "Visual detection of novel terrain via two-class classification," in *Proc. 2009 ACM Symposium on Applied Computing*, pp. 1145–1150 (2009).
10. I. Halatci, C. A. Brooks, and K. Iagnemma, "Terrain classification and classifier fusion for planetary exploration rovers," in *IEEE Aerospace Conf.*, pp. 1–11 (2007).
11. M. Happold, M. Ollis, and N. Johnson, "Enhancing supervised terrain classification with predictive unsupervised learning," in *Proc. Robotics: Science and Systems*, pp. 112–119 (2006).
12. D. Sadhukhan, "Autonomous ground vehicle terrain classification using internal sensors," Master's Thesis, Florida State University, Tallahassee, FL (2004).
13. J. Xu and S. Yue, "Mimicking visual searching with integrated top down cues and low-level features," *Neurocomputing* **133**, 1–17 (2014).
14. J. Allman and F. Miezin, "Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons," *Ann. Rev. Neurosci.* **8**, 407–430 (1985).

15. G. S. Cymbalyuk et al., "Modeling alternation to synchrony with inhibitory coupling: a neuromorphic VLSI approach," *Neural Comput.* **12**(10), 2259–2278 (2000).
16. G. Indiveri, "A current-mode hysteretic winner-take-all network, with excitatory and inhibitory coupling," *Analog Integr. Circuits Signal Process.* **28**(3), 279–29 (2001).
17. A. Destexhe et al., "Are corticothalamic 'up' states fragments of wakefulness?," *Trends Neurosci.* **30**(7), 334–342 (2007).
18. H. Heinze et al., "Visual event-related potentials index focused attention within bilateral arrays. I. Evidence for early selection," *Electroencephalograph. Clin. Neurophysiol.* **75**, 511–527 (1990).
19. L. Ojeda et al., "Terrain characterization and classification with a mobile robot," *J. Field Rob.* **23**(2), 103–122 (2006).
20. K. W. Chang, M. Ringgaard, and C. J. Lin, "Training and testing low-degree polynomial data mappings via linear SVM," *J. Mach. Learn. Res.* **11**, 1471–1490 (2010).
21. L. Bruzzone, D. Prieto, and S. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.* **37**(3), 1350–1359 (1999).
22. J. Inglada, "Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features," *ISPRS J. Photogramm. Remote Sens.* **62**(3), 236–248 (2007).
23. G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: a review," *ISPRS J. Photogramm. Remote Sens.* **66**(3), 247–259 (2011).
24. C. Weiss, H. Fröhlich, and A. Zell, "Vibration-based terrain classification using support vector machines," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems* (2006).
25. L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998).
26. Y. Xua and P. Duygulub, "Object-based image labeling through learning by example and multi-level segmentation," *Pattern Recognit.* **36**(6), 1407–1423 (2003).
27. http://emerald.ucsc.edu/~hyperwww/instruments.html.
28. http://peipa.essex.ac.uk/benchmark/databases/.
29. http://www.remotesensing.csdb.cn/.
30. B. Tso and R. C. Olsen, "Combining spectral and spatial information into hidden Markov models for unsupervised image classification," *Int. J. Remote Sens.* **26**(10), 2113–2133 (2005).
31. J. Zhang and G. M. Foody, "A fuzzy classification of sub-urban land cover from remotely sensed imagery," *Int. J. Remote Sens.* **19**(14), 2721–2738 (1998).
32. W. A. Harvey, "Performance evaluation for road extraction," *Bulletin de la Societe Francaise de Photogrammetrie et Teledetection* **153**(153), 79–97 (1999).
33. L. M. Keeping and P. E. Levy, "Performance appraisal reactions: measurement, modeling, and method bias," *J. Appl. Psychol.* **85**(5), 708–723 (2000).
34. J. Yang et al., "A modified Gabor filter design method for fingerprint image enhancement," *Pattern Recognit. Lett.* **24**(12), 1805–1817 (2003).
35. E. Arias-Castro and D. L. Donoho, "Does median filtering truly preserve edges better than linear filtering?," *Ann. Stat.* **37**(3), 1172–2009 (2009).
36. J. Xu, S. Yue, and Y. Tang, "A motion attention model based on rarity weighting and motion cues in dynamic scenes," *Int. J. Pattern Recognit. Artif. Intell.* **27**(6) (2013).
37. S. Yue and R. F. Claire, "Redundant neural vision systems–competing for collision recognition roles," *IEEE Trans. Auton. Mental Dev.* **5**(2), 173–186 (2013).

Jiawei Xu received his BS and MS degrees in computer engineering from Shanghai University of Engineering Science and Technology, Shanghai, China, 2007, and Hallym University, Republic of Korea, 2010, respectively. Now he is pursuing a PhD in the School of Computer Science, University of Lincoln, United Kingdom. His research interests include computer vision, human attention models, and visual cortex modeling. He was awarded the Outstanding Paper Award in Korea Multimedia Society (KMMS), 2010.

Ruisheng Wang joined the Department of Geomatics Engineering at the University of Calgary as an assistant professor in 2012. Prior to that, he worked as an industrial researcher at the Nokia

Location & Commerce (formerly NAVTEQ) in Chicago, USA, since 2008. His primary research focus was on mobile LiDAR data processing for next generation map making and navigation technology. He holds a PhD in electrical and computer engineering from McGill University.

Shigang Yue is a professor of computer science in the Lincoln School of Computer Science, University of Lincoln, United Kingdom. His research interests are mainly within the field of artificial intelligence, computer vision, robotics, brains, and neuroscience. He is particularly interested in biological visual neural systems, evolution of neuronal subsystems and their applications, e.g., in collision detection for vehicles, interactive systems, and robotics.